ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/494,297

DATE: 04/16/2001
TIME: 17:21:52

Input Set : A:\SequenceListing.09-494297.txt
Output Set: N:\CRF3\04162001\I494297.raw

```
 3 <110> APPLICANT: PODBIELSKI, ANDREAS
 5 <120> TITLE OF INVENTION: COLLAGEN-BINDING PROTEINS FROM STREPTOCOCCUS PYOGENES
 7 <130> FILE REFERENCE: P06628US0/BAS
 9 <140> CURRENT APPLICATION NUMBER: 09/494,297
10 <141> CURRENT FILING DATE: 2000-01-31
12 <160> NUMBER OF SEQ ID NOS: 4
14 <170> SOFTWARE: PatentIn Ver. 2.0
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 2274
18 <212> TYPE: DNA
19 <213> ORGANISM: Streptococcus pyogenes
21 <400> SEQUENCE: 1
22 atgaaaaaaa caaggtttcc aaataagctt aatactctta atactcaaag ggtattaagt 60
23 aaaaactcaa aacgatttac tgtcacttta gtgggagtct ttttaatgat cttcgctttg 120
24 gtaacttcca tggttggtgc taagactgtt tttggtttag tagaatcctc gacgccaaac 180
25 gcaataaatc cagattcaag ttcggaatac agatggtatg gatatgaatc ttatgtaaga 240
26 gggcatccat attataaaca gtttagagta gcacacgatt taagggttaa cttagaagga 300
27 agtagaagtt atcaagttta ttgctttaat ttaaagaaag catttcctct cggatcagat 360
28 agtagtgtta aaaagtggta taaaaaacat gatggaatct ctacaaaatt tgaagattat 420
29 gcgatgagcc ctagaattac gggagatgag ctaaatcaga agttacgagc tgttatgtat 480
30 aatggacatc cacaaaatgc caatggtatt atggaaggct ggaacccctt gaatgctatc 540
31 agagttacac aagaggcggt atggtactat tctgataatg ctcctatttc taatccagat 600
32 gaaagtttta aaagggagtc agaaagtaac ttggttagta cttctcaatt atctttgatg 660
33 cgtcaagctt tgaagcaact gattgatccg aatttggcaa ctaaaatgcc aaaacaagtt 720
34 ccggatgatt ttcagctaag tattttttgag tctgaggaca agggagataa atataataaa 780
35 ggataccaaa atcttttgag tggtggttta gttcctacta aaccaccaac tccaggagac 840
36 ccaccaatgc ctccaaatca acctcaaacg acttcagtac ttattagaaa gtatgctata 900
37 ggtgattact ctaaattgct tgaaggtgca acattacagt tgacagggga taacgtgaat 960
38 agttttcaag cgagagtgtt tagcagtaat gatattggag aaagaattga actatcagat 1020
39 ggaacttata ctttaactga attgaattct ccagctggtt atagtatcgc agagccaatc 1080
40 acttttaagg ttgaagctgg caaagtgtat actattattg atggaaaaca gattgaaaat 1140
41 cccaataaag agatagtaga gccttactca gtagaagcat ataatgattt tgaagaattt 1200
42 agcgttttaa ctacacaaaa ctatgcaaaa ttttattatg caaaaaataa aaatggaagt 1260
43 tcacaggttg tctattgctt taatgcagat ctaaaatctc accagactc tgaagatggt 1320
44 gggaaaacaa tgactccaga ctttacaaca ggagaagtaa aatacactca tattgcaggt 1380
45 cgtgacctct ttaaatatac tgtgaaacca agagataccg atcctgacac tttcttaaaa 1440
46 catatcaaaa aagtaattga gaagggttac agggaaaaag gacaagctat tgagtatagt 1500
47 ggtctaactg agacacaatt gcgtgcggct actcagttag caatatatta tttcactgat 1560
48 agtgctgaat tagataagga taaactaaaa gactatcatg gttttggaga catgaatgat 1620
49 agtactttag cagttgctaa aatccttgta gaatacgctc aagatagtaa tcctccacag 1680
50 ctaactgacc ttgatttctt tattccgaat aacaataaat atcaatctct tattggaact 1740
51 cagtggcatc cagaagattt agttgatatt attcgtatgg aagataaaaa agaagttata 1800
52 cctgtaactc ataatttaac attgagaaaa acggtgactg gtttagctgg tgacagaact 1860
53 aaagatttcc attttgaaat tgaattaaaa aataataagc aagaattgct ttctcaaact 1920
54 gttaaaacag ataaaacaaa cctcgaattt aaagatggta agcaaccat taatttaaaa 1980
55 catggggaaa gtttaacact tcaaggttta ccagaaggtt attcttacct tgtcaaagaa 2040
56 acagattctg aaggctataa ggttaaagtt aatagccaag aagtagcaaa tgctacagtt 2100
```

RAW SEQUENCE LISTING   DATE: 04/16/2001
PATENT APPLICATION: US/09/494,297   TIME: 17:21:52

Input Set : A:\SequenceListing.09-494297.txt
Output Set: N:\CRF3\04162001\I494297.raw

```
57 tcaaaaacag gaataacaag tgatgagaca cttgcttttg aaaataataa agagcctgtt 2160
58 gttcctacag gagttgatca aaagatcaat ggctatctag ctttgatagt tatcgctggt 2220
59 atcagtttgg ggatctgggg aattcacacg ataaggataa gaaaacatga ctag      2274
61 <210> SEQ ID NO: 2
62 <211> LENGTH: 757
63 <212> TYPE: PRT
64 <213> ORGANISM: Streptococcus pyogenes
66 <400> SEQUENCE: 2
67 Met Lys Lys Thr Arg Phe Pro Asn Lys Leu Asn Thr Leu Asn Thr Gln
68   1               5                  10                  15
70 Arg Val Leu Ser Lys Asn Ser Lys Arg Phe Thr Val Thr Leu Val Gly
71              20                  25                  30
73 Val Phe Leu Met Ile Phe Ala Leu Val Thr Ser Met Val Gly Ala Lys
74          35                  40                  45
76 Thr Val Phe Gly Leu Val Glu Ser Ser Thr Pro Asn Ala Ile Asn Pro
77      50                  55                  60
79 Asp Ser Ser Ser Glu Tyr Arg Trp Tyr Gly Tyr Glu Ser Tyr Val Arg
80   65                  70                  75                  80
82 Gly His Pro Tyr Tyr Lys Gln Phe Arg Val Ala His Asp Leu Arg Val
83                  85                  90                  95
85 Asn Leu Glu Gly Ser Arg Ser Tyr Gln Val Tyr Cys Phe Asn Leu Lys
86              100                 105                 110
88 Lys Ala Phe Pro Leu Gly Ser Asp Ser Ser Val Lys Lys Trp Tyr Lys
89              115                 120                 125
91 Lys His Asp Gly Ile Ser Thr Lys Phe Glu Asp Tyr Ala Met Ser Pro
92          130                 135                 140
94 Arg Ile Thr Gly Asp Glu Leu Asn Gln Lys Leu Arg Ala Val Met Tyr
95 145                 150                 155                 160
97 Asn Gly His Pro Gln Asn Ala Asn Gly Ile Met Glu Gly Leu Glu Pro
98                  165                 170                 175
100 Leu Asn Ala Ile Arg Val Thr Gln Glu Ala Val Trp Tyr Tyr Ser Asp
101             180                 185                 190
103 Asn Ala Pro Ile Ser Asn Pro Asp Glu Ser Phe Lys Arg Glu Ser Glu
104             195                 200                 205
106 Ser Asn Leu Val Ser Thr Ser Gln Leu Ser Leu Met Arg Gln Ala Leu
107     210                 215                 220
109 Lys Gln Leu Ile Asp Pro Asn Leu Ala Thr Lys Met Pro Lys Gln Val
110 225                 230                 235                 240
112 Pro Asp Asp Phe Gln Leu Ser Ile Phe Glu Ser Glu Asp Lys Gly Asp
113             245                 250                 255
115 Lys Tyr Asn Lys Gly Tyr Gln Asn Leu Leu Ser Gly Gly Leu Val Pro
116             260                 265                 270
118 Thr Lys Pro Pro Thr Pro Gly Asp Pro Pro Met Pro Pro Asn Gln Pro
119     275                 280                 285
121 Gln Thr Thr Ser Val Leu Ile Arg Lys Tyr Ala Ile Gly Asp Tyr Ser
122     290                 295                 300
124 Lys Leu Leu Glu Gly Ala Thr Leu Gln Leu Thr Gly Asp Asn Val Asn
125 305                 310                 315                 320
127 Ser Phe Gln Ala Arg Val Phe Ser Ser Asn Asp Ile Gly Glu Arg Ile
```

```
128                 325                 330                 335
130 Glu Leu Ser Asp Gly Thr Tyr Thr Leu Thr Glu Leu Asn Ser Pro Ala
131             340                 345                 350
133 Gly Tyr Ser Ile Ala Glu Pro Ile Thr Phe Lys Val Glu Ala Gly Lys
134             355                 360                 365
136 Val Tyr Thr Ile Ile Asp Gly Lys Gln Ile Glu Asn Pro Asn Lys Glu
137         370                 375                 380
139 Ile Val Glu Pro Tyr Ser Val Glu Ala Tyr Asn Asp Phe Glu Glu Phe
140 385                 390                 395                 400
142 Ser Val Leu Thr Thr Gln Asn Tyr Ala Lys Phe Tyr Tyr Ala Lys Asn
143             405                 410                 415
145 Lys Asn Gly Ser Ser Gln Val Val Tyr Cys Phe Asn Ala Asp Leu Lys
146             420                 425                 430
148 Ser Pro Pro Asp Ser Glu Asp Gly Gly Lys Thr Met Thr Pro Asp Phe
149         435                 440                 445
151 Thr Thr Gly Glu Val Lys Tyr Thr His Ile Ala Gly Arg Asp Leu Phe
152     450                 455                 460
154 Lys Tyr Thr Val Lys Pro Arg Asp Thr Asp Pro Asp Thr Phe Leu Lys
155 465                 470                 475                 480
157 His Ile Lys Lys Val Ile Glu Lys Gly Tyr Arg Glu Lys Gly Gln Ala
158             485                 490                 495
160 Ile Glu Tyr Ser Gly Leu Thr Glu Thr Gln Leu Arg Ala Ala Thr Gln
161             500                 505                 510
163 Leu Ala Ile Tyr Tyr Phe Thr Asp Ser Ala Glu Leu Asp Lys Asp Lys
164         515                 520                 525
166 Leu Lys Asp Tyr His Gly Phe Gly Asp Met Asn Asp Ser Thr Leu Ala
167     530                 535                 540
169 Val Ala Lys Ile Leu Val Glu Tyr Ala Gln Asp Ser Asn Pro Pro Gln
170 545                 550                 555                 560
172 Leu Thr Asp Leu Asp Phe Phe Ile Pro Asn Asn Asn Lys Tyr Gln Ser
173             565                 570                 575
175 Leu Ile Gly Thr Gln Trp His Pro Glu Asp Leu Val Asp Ile Ile Arg
176         580                 585                 590
178 Met Glu Asp Lys Lys Glu Val Ile Pro Val Thr His Asn Leu Thr Leu
179         595                 600                 605
181 Arg Lys Thr Val Thr Gly Leu Ala Gly Asp Arg Thr Lys Asp Phe His
182     610                 615                 620
184 Phe Glu Ile Glu Leu Lys Asn Asn Lys Gln Glu Leu Leu Ser Gln Thr
185 625                 630                 635                 640
187 Val Lys Thr Asp Lys Thr Asn Leu Glu Phe Lys Asp Gly Lys Ala Thr
188             645                 650                 655
190 Ile Asn Leu Lys His Gly Glu Ser Leu Thr Leu Gln Gly Leu Pro Glu
191         660                 665                 670
193 Gly Tyr Ser Tyr Leu Val Lys Glu Thr Asp Ser Glu Gly Tyr Lys Val
194     675                 680                 685
196 Lys Val Asn Ser Gln Glu Val Ala Asn Ala Thr Val Ser Lys Thr Gly
197     690                 695                 700
199 Ile Thr Ser Asp Glu Thr Leu Ala Phe Glu Asn Asn Lys Glu Pro Val
200 705                 710                 715                 720
```

```
                  RAW SEQUENCE LISTING              DATE: 04/16/2001
                  PATENT APPLICATION:  US/09/494,297  TIME: 17:21:53

                  Input Set : A:\SequenceListing.09-494297.txt
                  Output Set: N:\CRF3\04162001\I494297.raw
```

```
202 Val Pro Thr Gly Val Asp Gln Lys Ile Asn Gly Tyr Leu Ala Leu Ile
203             725             730                 735
205 Val Ile Ala Gly Ile Ser Leu Gly Ile Trp Gly Ile His Thr Ile Arg
206             740             745                 750
208 Ile Arg Lys His Asp
209         755
212 <210> SEQ ID NO: 3
213 <211> LENGTH: 2229
214 <212> TYPE: DNA
215 <213> ORGANISM: Streptococcus pyogenes
217 <400> SEQUENCE: 3
218 ttgcaaaaga gggataaaac caattatgga agcgctaaca acaaacgacg acaaacgacg 60
219 atcggattac tgaaagtatt tttgacgttt gtagctctga taggaatagt agggtttttct 120
220 atcagagcgt tcggagctga agaacaatca gtaccaaata gacaaagctc aattcaagat 180
221 tatccgtggt atggctatga ttcttatcct aaaggctacc cagactatag tccgttaaag 240
222 acttaccata atttaaaagt aaatttagag ggaagtaagg attatcaagc atactgcttt 300
223 aatttaacaa aacattttcc atccaagtca gatagtgtta gatcacaatg gtataaaaaa 360
224 cttgaaggaa ctaatgaaaa ctttatcaag ttagcagata aaccaagaat agaagacgga 420
225 cagttacaac aaaatatatt gaggattctc tataatggat atcctaataa tcgtaatggg 480
226 ataatgaaag ggatagatcc tctaaacgct attttagtga ctcaaaatgc tatttggtat 540
227 actgattcag ctcaaattaa tccggatgaa agttttaaaa cagaagctcg aagtaatggt 600
228 attaatgacc agcagttagg cttaatgcga aaagctttaa aagaactaat tgatccaaac 660
229 ttagggtcaa aatattcgaa taaaactcca tcaggttatc ggttaaatgt atttgaatct 720
230 catgataagc ctttccaaaa tctttttgagt gctgagtatg ttccggatac tccccccaaaa 780
231 ccaggagaag agcctccggc taaaactgaa aaaacatcag tcattatcag aaaatatgcg 840
232 gaaggtgact ctaaacttct agagggagca accttaaagc tttctcaaat tgaaggaagt 900
233 ggttttcaag aaaaagactt tcaaagtaat agtttaggag aaactgtcga attaccaaat 960
234 gggacttata ccttaacaga aacatcatct ccagatggat ataaaattgc ggagccgatt 1020
235 aagtttagag tagagaataa aaaagtattt atcgtccaaa aagatggttc tcaagtggaa 1080
236 aatccaaaca aagaagtagc agagccatac tcagtggaag cgtataatga ctttatggat 1140
237 gaagaagtac tctcgggttt tactccatac ggaaaattct attacgctac aaataaggat 1200
238 aaaagttcac aagttgtcta ctgcttcaat gctgatttac actcaccacc tgactcatat 1260
239 gatagtggtg agactataaa tccagatact agtacgatga aagaagtcaa gtacacacat 1320
240 acggcaggta gtgacttgtt taaatatgcg ctaagaccga gagatacaaa tccagaagac 1380
241 ttcttaaagc acattaaaaa agtaattgaa aaaggctaca agaaaaaagg tgatagctat 1440
242 aatggattaa cagaaacaca gtttcgcgcg gctactcagc ttgctatcta ttatttttaca 1500
243 gacagtgctg acttaaaaac cttaaaaact tataacaatg ggaaaggtta ccatggatttt 1560
244 gaatctatgg atgaaaaaac cctagctgtc acaaaagaat taattactta tgctcaaaat 1620
245 ggcagtgccc ctcaactaac aaatcttgat ttcttcgtac ctaataatag caaagaccaa 1680
246 tctcttattg ggacagaatg ccatccagat gatttggttg acgtgattcg tatggaagat 1740
247 aaaaagcaag aagttattcc agtaactcac agtttgacag tgaaaaaaac agtagtcggt 1800
248 gagttgggag ataaaactaa aggcttccaa tttgaacttg agttgaaaga taaaactgga 1860
249 cagcctattg ttaacactct aaaaactaat aatcaagatt tagtagctaa agatgggaaa 1920
250 tattcatta atctaaagca tggtgacacc ataagaatag aaggattacc gacgggatat 1980
251 tcttatactc tgaaagaggc tgaagctaag gattatatag taaccgttga taacaaagtt 2040
252 agtcaagaag cgcagtcagt aggtaaggat ataacagaag acaaaaaagt cactttttgaa 2100
253 aaccgaaaag atcttgtccc accaactggt ttgacaacag atggggctat ctatctttgg 2160
254 ttgttattac ttgttccact tgggttattg gtttggctat ttggtcgtaa agggttaaaa 2220
255 aatgactaa                                          2229
```

RAW SEQUENCE LISTING   DATE: 04/16/2001
PATENT APPLICATION: US/09/494,297   TIME: 17:21:53

Input Set : A:\SequenceListing.09-494297.txt
Output Set: N:\CRF3\04162001\I494297.raw

```
257 <210> SEQ ID NO: 4
258 <211> LENGTH: 742
259 <212> TYPE: PRT
260 <213> ORGANISM: Streptococcus pyogenes
262 <400> SEQUENCE: 4
263 Met Gln Lys Arg Asp Lys Thr Asn Tyr Gly Ser Ala Asn Asn Lys Arg
264  1               5                  10                  15
266 Arg Gln Thr Thr Ile Gly Leu Leu Lys Val Phe Leu Thr Phe Val Ala
267             20                  25                  30
269 Leu Ile Gly Ile Val Gly Phe Ser Ile Arg Ala Phe Gly Ala Glu Glu
270             35                  40                  45
272 Gln Ser Val Pro Asn Arg Gln Ser Ser Ile Gln Asp Tyr Pro Trp Tyr
273     50                  55                  60
275 Gly Tyr Asp Ser Tyr Pro Lys Gly Tyr Pro Asp Tyr Ser Pro Leu Lys
276  65                  70                  75                  80
278 Thr Tyr His Asn Leu Lys Val Asn Leu Glu Gly Ser Lys Asp Tyr Gln
279                 85                  90                  95
281 Ala Tyr Cys Phe Asn Leu Thr Lys His Phe Pro Ser Lys Ser Asp Ser
282             100                 105                 110
284 Val Arg Ser Gln Trp Tyr Lys Lys Leu Glu Gly Thr Asn Glu Asn Phe
285             115                 120                 125
287 Ile Lys Leu Ala Asp Lys Pro Arg Ile Glu Asp Gly Gln Leu Gln Gln
288     130                 135                 140
290 Asn Ile Leu Arg Ile Leu Tyr Asn Gly Tyr Pro Asn Asn Arg Asn Gly
291 145                 150                 155                 160
293 Ile Met Lys Gly Ile Asp Pro Leu Asn Ala Ile Leu Val Thr Gln Asn
294             165                 170                 175
296 Ala Ile Trp Tyr Thr Asp Ser Ala Gln Ile Asn Pro Asp Glu Ser Phe
297             180                 185                 190
299 Lys Thr Glu Ala Arg Ser Asn Gly Ile Asn Asp Gln Gln Leu Gly Leu
300             195                 200                 205
302 Met Arg Lys Ala Leu Lys Glu Leu Ile Asp Pro Asn Leu Gly Ser Lys
303     210                 215                 220
305 Tyr Ser Asn Lys Thr Pro Ser Gly Tyr Arg Leu Asn Val Phe Glu Ser
306 225                 230                 235                 240
308 His Asp Lys Pro Phe Gln Asn Leu Leu Ser Ala Glu Tyr Val Pro Asp
309             245                 250                 255
311 Thr Pro Pro Lys Pro Gly Glu Glu Pro Pro Ala Lys Thr Glu Lys Thr
312             260                 265                 270
314 Ser Val Ile Ile Arg Lys Tyr Ala Glu Gly Asp Ser Lys Leu Leu Glu
315             275                 280                 285
317 Gly Ala Thr Leu Lys Leu Ser Gln Ile Glu Gly Ser Gly Phe Gln Glu
318     290                 295                 300
320 Lys Asp Phe Gln Ser Asn Ser Leu Gly Glu Thr Val Glu Leu Pro Asn
321 305                 310                 315                 320
323 Gly Thr Tyr Thr Leu Thr Glu Thr Ser Ser Pro Asp Gly Tyr Lys Ile
324             325                 330                 335
326 Ala Glu Pro Ile Lys Phe Arg Val Glu Asn Lys Lys Val Phe Ile Val
327             340                 345                 350
```

VERIFICATION SUMMARY   DATE: 04/16/2001
PATENT APPLICATION: US/09/494,297   TIME: 17:21:54

Input Set : **A:\SequenceListing.09-494297.txt**
Output Set: **N:\CRF3\04162001\I494297.raw**